

Un modèle d'analyse de généalogie ascendante A Model for Analysing Ascending Genealogies Un modelo de análisis de genealogía ascendente

Esther Létourneau et Francine M. Mayer

Volume 17, numéro 2, automne 1988

URI : <https://id.erudit.org/iderudit/600640ar>

DOI : <https://doi.org/10.7202/600640ar>

[Aller au sommaire du numéro](#)

Éditeur(s)

Association des démographes du Québec

ISSN

0380-1721 (imprimé)

1705-1495 (numérique)

[Découvrir la revue](#)

Citer cet article

Létourneau, E. & Mayer, F. M. (1988). Un modèle d'analyse de généalogie ascendante. *Cahiers québécois de démographie*, 17(2), 213–232.
<https://doi.org/10.7202/600640ar>

Résumé de l'article

Cet article présente un modèle de généalogie ascendante, la définition et l'évaluation des fonctions qui permettent d'en faire l'étude, ainsi que les principaux algorithmes développés pour en faire des outils de travail sur micro-ordinateur. Les exemples de fonctions d'extraction de résultats destinés à l'analyse d'ascendances illustrent la portée de la mise au point du modèle, que ce soit pour la sortie de listes ou le calcul des coefficients de consanguinité et de complétude.

Un modèle d'analyse de généalogie ascendante

Esther LÉTOURNEAU et Francine M. MAYER*

INTRODUCTION

La parenté entre la démographie et la génétique des populations ne date pas d'hier. Mais si depuis l'après-guerre jusqu'à dans les années 1980 les développements multidisciplinaires sont restés peu nombreux, on ne peut manquer de souligner la récente reprise des travaux requérant simultanément l'expertise de démographes, de généticiens et d'historiens. La réalisation d'analyses diverses à partir de données structurées en généalogies ascendantes témoigne d'un renouveau au sein des préoccupations de recherche portant sur la dynamique des populations humaines.

Les travaux de Bideau et Poulain (1984), Bouchard et alii (1984, 1988), Mayer et alii (1984), Declos (1985) et Létourneau (1986) sont des exemples d'utilisation de généalogies ascendantes, en réponse à des problématiques spécifiques. Bideau et Poulain étaient préoccupés de démographie classique alors que les autres auteurs cherchaient à comprendre, à l'échelle d'une population, la fréquence élevée de certaines maladies héréditaires.

Ainsi, Bideau et Poulain ont calculé des indices classiques à partir des caractéristiques des individus apparaissant dans des généalogies ascendantes et ayant été sélectionnés par

* Équipe de recherches sur la dynamique des populations humaines (EDYPH), Département de Sciences biologiques, Université du Québec à Montréal.

La rédaction de cet article a été rendue possible grâce à des subventions du Conseil de recherches en sciences humaines du Canada (CRSH, no : 410-85-0404, 1985/1987), du Fonds pour la Formation de chercheurs et d'aide à la recherche (FCAR, no : EQ-2605, 1984/1987) et du Comité d'attribution de fonds internes de recherches de l'Université de Montréal (CAFIR, no : 167-1983).

échantillonnage. Quant à F.M. Mayer, elle s'attache, en regroupant les ancêtres fondateurs qui constituent les fins d'information généalogique, à démontrer la persistance d'une double endogamie dans la population de l'île de Saint-Barthélemy. Dans son étude, E. Létourneau compare les structures généalogiques et calcule les indices de nuptialité des couples d'ancêtres inscrits dans des généalogies reconstituées pour la même population. Les travaux de Bouchard et alii et de Declos, concentrés sur la population du Saguenay, utilisent des ascendances d'une part pour identifier et caractériser les fondateurs (nombre et origine géographique) et d'autre part pour calculer l'apparentement et la consanguinité. De surcroît, ils présentent des indices démographiques des groupes de couples sélectionnés à même les généalogies.

Bien que le traitement des généalogies ascendantes ne soit pas uniforme d'un auteur à l'autre, il existe bien une similitude des méthodologies employées. Ainsi, une généalogie ascendante représente un ensemble de couples pouvant, en tout ou en partie, faire l'objet d'analyses démographiques ou génétiques. Toute généalogie ascendante constitue une unité en soi pouvant se caractériser par des paramètres précis, comme le coefficient de consanguinité qui permet d'évaluer, pour l'individu en tête de généalogie, la probabilité qu'il ait reçu deux copies d'un même gène provenant d'un ancêtre commun à ses parents. Par ailleurs, correspondant à une structure formée d'individus, chaque généalogie peut se voir attribuer des caractéristiques de complétude, suivant le nombre d'individus qui la composent et suivant le nombre de positions que ceux-ci y occupent (nombre de répétitions d'un ancêtre dans une généalogie ascendante).

Dans un contexte de développement scientifique, il importe de définir les méthodologies qui permettent l'obtention des résultats attendus. Un modèle a ainsi été mis au point pour le traitement de généalogies ascendantes. Parce qu'il permet la transposition de toute généalogie en des paramètres fixes, ce modèle constitue un premier effort de systématisation et de définition de l'approche généalogique. Son efficacité a été testée lors de l'intégration à OGAIP, un logiciel de gestion et d'analyse de registres de population structurés en base de données, qui a été développé sur micro-ordinateur¹. Le module

1. Le développement d'un «outil de gestion et d'analyse d'isolats de population» (OGAIP) initié par EDYPH, sous la supervision du Centre de calcul de l'Université de Montréal, se poursuit avec la collaboration du PRDH (Département de démographie, Université de Montréal) et du Groupe R.O. (Université de Lyon II).

d'analyses génétiques d'OGAIP a été élaboré à partir de ce modèle, et il regroupe les fonctions définies pour fournir automatiquement des résultats tels que ceux mentionnés.

La présentation du modèle de généalogie ascendante, la définition et l'évaluation des fonctions qui permettent d'en faire l'étude, ainsi que les principes algorithmiques développés pour en tirer des outils de travail sur micro-ordinateur, font l'objet de cet article.

DÉFINITION D'UNE GÉNÉALOGIE ASCENDANTE

Une généalogie ascendante correspond à une unité définie par un ensemble d'objets organisés de façon logique. Ainsi, une généalogie ascendante est un ensemble fini et structuré d'individus qui, tous, sont les ancêtres à différentes générations d'un individu distinct des précédents et que l'on appelle ego. Ego constitue le point initial alors que la fin d'une généalogie ascendante est indiquée par les ancêtres fondateurs, individus dont les parents sont inconnus, et par les semi-fondateurs, individus dont un seul des parents est inconnu (voir figure 1). Une numérotation des positions possibles où se situent les ancêtres, empruntée à Sosa (Pélissier, 1983), reproduit la charpente de toute ascendance (voir figure 2). Ainsi, lors de la reconstitution automatique, ou même manuelle, d'une généalogie, chaque ancêtre est associé à un rang (nombre entier) qui permet de le localiser par rapport à ego. Précisons qu'un individu est associé à autant de rangs qu'on lui compte de mentions comme ancêtre à l'intérieur d'une même généalogie.

DE LA DÉFINITION AU MODÈLE GÉNÉALOGIQUE INFORMATISÉ

Notre modèle généalogique a été conçu dans le cadre du développement d'un logiciel de gestion et d'analyses sur micro-ordinateur du registre de population de l'île Saint-Barthélemy. Les outils informatiques utilisés pour la construction du logiciel OGAIP, soit le langage Prolog et plus spécialement le progiciel DBase III+, ont servi à tester l'efficacité du modèle.

Figure 1
Éléments d'une généalogie

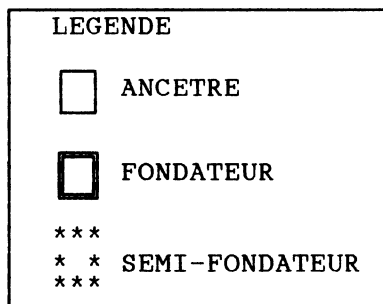
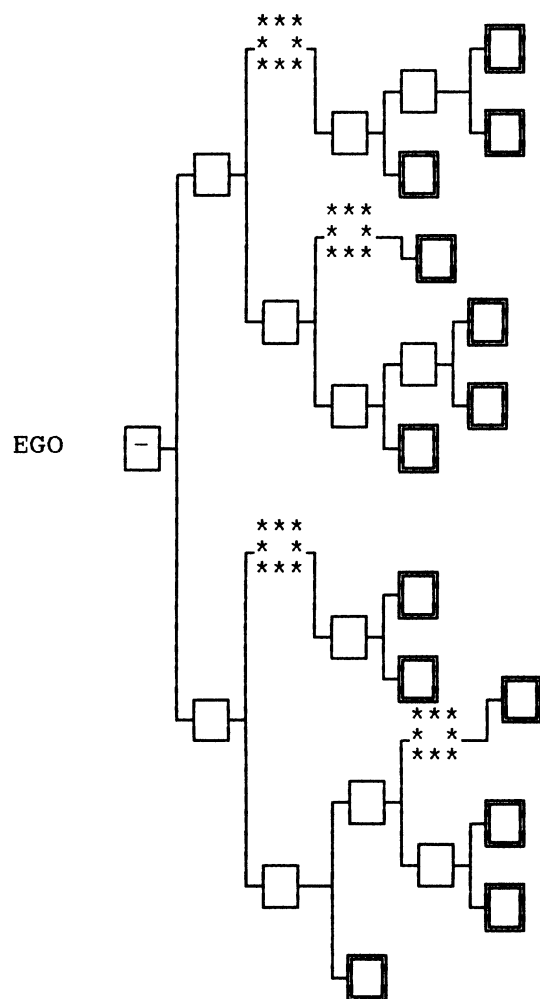
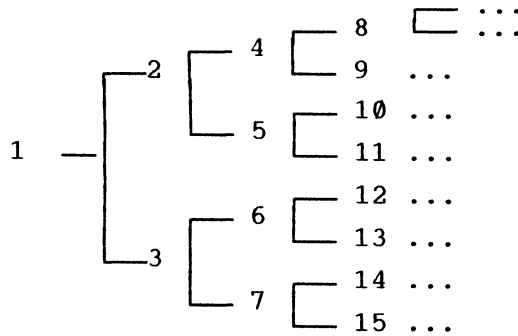


Figure 2
Ordre des rangs dans une généalogie ascendante



La reconstitution d'une généalogie repose sur l'exploitation d'un fichier qui réunit tous les individus inscrits au registre. Ce fichier initial qui est la base du registre de population informatisé comprend, entre autres, une variable numérique (matricule) constituant l'identificateur propre à chacun. En inscrivant dans l'enregistrement d'un individu, outre son matricule, celui de son père et celui de sa mère, et ce, de façon systématique pour l'ensemble des individus du fichier, on ouvre la voie à une lecture en chaîne pouvant conduire à la création d'une généalogie ascendante.

Cette étape du passage de la définition d'une généalogie à sa reconstitution automatique a exigé la conception d'un algorithme effectuant une lecture récursive des triplets «individu-père-mère» et attribuant à chaque ancêtre identifié, le ou les rangs qu'il occupe par rapport à ego. Cet algorithme amène, pour la construction d'une seule généalogie, la création et le traitement de plusieurs fichiers auxiliaires. Ce sont ces fichiers qui permettent de créer et de rassembler les données qui, au terme de l'opération algorithmique, constituent les éléments propres à chaque généalogie ascendante. Le tableau 1 présente la structure du fichier individus qui est à l'origine

Tableau 1

Structure des fichiers individus et généalogie ascendante

A. INDIVIDUS

Variable	Type	Signification
IDI	Num.	Matricule de l'individu
PÈRE	Num.	Matricule de son père
MÈRE	Num.	Matricule de sa mère
....	Variables socio-démographiques
CNGCOEF	Num.	Coefficient de consanguinité de IDI

B. GÉNÉALOGIE ASCENDANTE

Variable	Type	Signification
IDI	NUM.	Matricule d'un individu de la généalogie ascendante
NGEN	NUM.	Génération correspondant au rang de IDI
SEXE	CAR.	Sexe de IDI
FLAGSORT ^a	CAR.	Transcription alphabétique du rang
RANG	NUM.	Rang de IDI dans la généalogie
CROSSREF ^a	NUM.	Le plus petit rang de IDI

a. Ces variables servent essentiellement au calcul des fonctions d'analyse; elles complètent les quatre autres variables qui, elles, servent spécifiquement à la reconstitution des généalogies.

de toute ascendance et celle du fichier généalogique final. Si le nombre d'enregistrements du premier correspond au nombre d'individus distincts du registre de population, l'organisation des données diffère pour le fichier des généalogies. Pour celui-ci, en effet, chaque ascendance correspond à un fichier indépendant; il existe donc autant de fichiers généalogiques qu'il y a d'ascendances à compléter. Précisons qu'on compte dans un fichier généalogique autant d'enregistrements que de rangs pour lesquels on a pu identifier l'ancêtre correspondant inscrit au registre de population.

L'obtention de résultats tels le coefficient de consanguinité, le coefficient de complétude d'une généalogie ou la liste des individus qui y apparaissent comme ancêtres, qu'ils soient fondateurs ou non, découle principalement de l'exploitation du fichier généalogique. Les algorithmes qui permettent la sortie automatique de ces valeurs sont considérés ici comme des fonctions d'analyse de généalogies ascendantes.

FONCTIONS D'ANALYSE DES GÉNÉALOGIES

Les fonctions d'analyse de généalogies ascendantes peuvent concerner tant les individus qui s'y inscrivent que l'unité généalogique comme telle.

1. Fonctions d'extraction de listes d'ancêtres

Les plus simples fonctions sont évidemment celles dont le résultat se résume à une liste d'individus extraits selon des critères préétablis : liste des fondateurs d'une généalogie ou d'un groupe de généalogies; liste des semi-fondateurs, liste des parents d'ego, etc.

Techniquement parlant, ces résultats s'obtiennent par comparaison des enregistrements relatifs à un même individu d'un fichier à l'autre. Sans pour autant entreprendre une description exhaustive du processus informatique requis, il importe de souligner que le mode de gestion du registre de population joue un rôle fondamental dans la conception des algorithmes de telles fonctions. Une organisation du registre en base de données relationnelles, comme on le retrouve dans OGAIP, prévoit spécifiquement le rapprochement des divers

fichiers à travers les enregistrements propres à chacun. Une simple commande du logiciel permet le rapprochement des enregistrements communs à plusieurs fichiers.

À titre d'exemple, reprenons la fonction d'extraction de listes et imaginons qu'un individu de matricule X fasse l'objet d'un rapprochement entre le fichier des individus et celui d'une généalogie ascendante. Sachant que dans le fichier des individus on identifie le statut généalogique (fondateur, semi-fondateur ou non-fondateur) par la présence de matricules dans les champs réservés aux parents, un test sur ces variables permet de déterminer si X est, ou non, un fondateur. Transposée à l'ensemble des individus inscrits dans la généalogie considérée, l'application de ce principe permet la sortie de la liste de tous les fondateurs.

Le lecteur comprendra qu'une fois la liste en main, il devient relativement simple de réaliser des calculs démographiques, ou d'autres manipulations, la difficulté de l'opération ne dépendant plus des embûches que réserve le traitement de généalogies complexes, mais bien du volume de la liste extraite ou de la qualité de l'information qui s'y rattache (dates inconnues ou de qualité douteuse, par exemple).

Toute liste peut être imprimée ou reproduite dans un fichier. Le tableau 2 présente un extrait d'une liste imprimée.

Tableau 2
Exemple de sortie d'une liste d'ancêtres

LISTE DES ANCÊTRES DE L'INDIVIDU # 3457

<u>Matricule</u>	<u>Nombre de mentions</u>	<u>Nom</u>	<u>Prénom</u>	<u>Rang</u>
1391	1	B.	Benoît	8
1392	1	O.	Geneviève	9
1399	1	B.	Gustave	4
1440	1	B.	Frédéric	2
1441	1	G.	Annemarie	3
3416	1	B.	Benoît	16
3417	2	L.	Suzanne	17, 25
...

Sources : Registre de population de l'île Saint-Barthélemy et logiciel OGAIP.

2. Fonctions d'analyse de complétude

On appelle analyse de complétude l'étude des structures d'une généalogie ascendante en termes de générations complétées et de rangs auxquels sont associés les ancêtres qui les composent. La complétude est un concept qui dépend directement de l'information que peut fournir le registre de population, parce que l'identification des ancêtres d'une généalogie est limitée à ceux qui sont inscrits dans cette banque de données.

La complétude a une valeur méthodologique fondamentale à deux points de vue. La signification du coefficient de consanguinité d'une généalogie, telle que la présente Jacquard (1974), fait nécessairement appel à ce préalable essentiel qu'est la connaissance de la complétude, et ce particulièrement dans le cadre d'études comparatives. En effet, plus on dispose d'informations généalogiques qui permettent de remonter loin dans l'ascendance d'un individu, plus on a de chances de trouver un ancêtre commun à son père et à sa mère, ce qui augmente la valeur de son coefficient de consanguinité. On réalise l'importance de la prise en compte de cette variable lorsqu'il s'agit de comparer, par exemple, la consanguinité moyenne de plusieurs populations. Comme la profondeur généalogique peut être très variable d'un groupe d'individus à l'autre, négliger ce facteur entraîne un biais si on compare des coefficients de consanguinité qui dépendent d'informations de nature plus ou moins complète. En ce sens, la complétude fait office de critique des sources des généalogies, parce qu'elle caractérise l'information permettant la reconstitution des généalogies.

Dans l'étude de Létourneau (1986), l'analyse de complétude est aussi un moyen d'évaluer la comparabilité de groupes de généalogies. Mais l'objectif est tout autre. Ainsi, après avoir observé que dans un corpus généalogique donné la profondeur généalogique de certaines ascendances était limitée dès les toutes premières générations, l'auteure cherche à identifier la cause de cette différence. Dépend-elle de lacunes dans les registres d'état civil utilisés pour dresser le registre de population à partir duquel les généalogies ont été reconstituées ? Ou faut-il plutôt chercher l'explication du côté du comportement des ancêtres concernés ?

L'intérêt de cette question est d'ordre critique. L'analyse du comportement démographique des couples sélectionnés au sein d'un corpus généalogique comportant des inégalités exige bien entendu l'évaluation des caractéristiques de la sous-population qu'il représente. L'exhaustivité des effectifs en constitue l'une des plus importantes. Envisagée de cette façon, la complétude apparaît comme une variable utile en démographie historique.

Tableau 3
Exemple des résultats de l'analyse de complétude

ANALYSE GLOBALE DE LA GÉNÉALOGIE ASCENDANTE DE L'INDIVIDU # 457

Nombre d'ancêtres répertoriés : 93

Nombre de fondateurs : 46

Nombre de semi-fondateurs : 1

ANALYSE GÉNÉRATIONNELLE

Profondeur : 7

Rangs occupés/maximum : 105/254

Complétude globale : 41,33 %

Génération	Rangs occ./max.	Complétude	Nombre de fondateurs	Cumul
1	2 / 2	100,00 %	0	0,00
2	4 / 4	100,00 %	0	0,00
3	8 / 8	100,00 %	0	0,00
4	16 / 16	100,00 %	1	0,25
5	30 / 32	93,75 %	13	2,28
6	33 / 64	51,56 %	27	4,81
7	12 / 128	9,38 %	12	5,47

COEFFICIENT DE COMPLÉTUDE

Moyenne de la profondeur généalogique : 5,47

Variance : 1,03

Sources : Registre de population de l'île Saint-Barthélemy et logiciel OGAIP.

Plusieurs fonctions font partie du volet complétude. Elles consistent d'abord en une analyse générationnelle, c'est-à-dire la prise en compte des rangs identifiés dans une généalogie par génération, en termes absolus et relatifs. S'y rajoute, tel qu'illustré au tableau 3, la répartition de ces rangs selon le nombre de fondateurs auxquels ils correspondent. La compilation de ces résultats est réalisée à partir de la variable génération du fichier généalogique, comme elle reprend en partie la procédure permettant d'identifier le statut des fondateurs.

Apparaissent également avec ces résultats, le coefficient moyen de complétude (moyenne pondérée des générations d'une généalogie) et la variance, deux paramètres qui synthétisent l'analyse générationnelle. Le calcul automatique de ces valeurs repose sur un algorithme qui, de nouveau, fait essentiellement intervenir une liste des fondateurs d'une généalogie et la génération où ils apparaissent, et qui est calqué sur la mesure d'entropie généalogique mise au point par Kouladjian (1986).

Cet auteur définit ainsi la valeur attendue de la génération de tous les fondateurs, valeur que nous désignons comme la moyenne de la profondeur généalogique :

$$S_b = -\sum_i P_i \log P_i$$

$$= \sum_i N_i / 2^{N_i}$$

où S_b = valeur attendue de la génération des fondateurs ou $\log 1/P_i$;

P_i = probabilité de l'origine du gène provenant du fondateur i de la généalogie R ;

\log = logarithme à base 2;

N_i = la génération du fondateur i .

La variance de cette moyenne se formule donc ainsi :

$$V = \sum P_i \log^2 P_i - S_b^2$$

Il importe toutefois de souligner que la moyenne et la variance demeurent des mesures approximatives, qui ne peuvent se substituer à la description exhaustive d'une généalogie offerte par la distribution par rangs. En effet, ces indices ne nous renseignent pas aussi adéquatement sur le recouvrement d'une généalogie que le fait une distribution selon le nombre de rangs identifiés par génération.

En dépit de l'intérêt des deux paramètres de complétude, il est donc essentiel de recourir aux distributions selon le rang, dans le cadre par exemple d'une analyse comparative entre généalogies ascendantes, notamment lorsque le coefficient de consanguinité constitue l'objet de comparaison.

3. Fonctions d'analyse de la consanguinité

Le coefficient de consanguinité sert à évaluer la probabilité qu'un individu soit homozygote pour un gène qu'il aurait reçu d'un ancêtre commun à ses parents, en descendance directe et sans mutation. Ce calcul est souvent utilisé pour comprendre la fréquence de pathologies héréditaires présentes dans certaines populations. De plus, le coefficient de consanguinité moyen F d'une population, obtenu par le rapport de la somme de tous les coefficients f de consanguinité des individus d'une génération donnée ou d'une période donnée (incluant les valeurs nulles) au nombre total d'individus, nous donne une valeur en probabilité du degré d'homogénéité génétique de cette population, c'est-à-dire de son appauvrissement en termes de diversité biologique.

À la différence de l'étude de la complétude, celle de la consanguinité représente depuis longtemps un enjeu de taille en termes de définition d'algorithme. Les essais de Kudo (1962) et de Garçon et Jacquard (1971) avaient bien permis d'obtenir des résultats valables, mais la structure de leur algorithme laissait place à une sous-estimation des coefficients, à cause d'un décompte incomplet des chaînes de consanguinité. Les développements récents de la micro-informatique et des langages de programmation ne sont pas sans avoir influencé la conception de l'algorithme implanté dans OGAIP, qui permet un calcul automatique du coefficient de consanguinité avec prise en compte de toutes les chaînes de consanguinité reliant entre eux, par un ancêtre commun, les père et mère d'ego².

Nous ne présenterons pas ici une description détaillée de cet algorithme. Nous croyons qu'il importe avant tout de dire qu'il a été élaboré à partir de la formule de Wright (1922), ainsi définie par Jacquard (1974) :

2. Cet algorithme a été développé par E. Heyer (Groupe R.O., Université de Lyon II).

$$f_i = \sum_{j=1}^k (0,5)^{m+p+1} (1 + f_{c_j})$$

- où f_i = le coefficient de consanguinité de la généalogie de l'ego i ;
 k = le nombre d'ancêtres communs aux parents d'ego (un même ancêtre commun qui apparaît plusieurs fois dans une même généalogie est recompté à chacune de ses apparitions);
 f_{c_j} = le coefficient de consanguinité de l'ancêtre commun j ;
 m = le nombre d'ancêtres qui séparent la mère d'ego de l'ancêtre commun (en ne comptant pas l'ancêtre commun);
 p = idem pour le père³.

Il est intéressant de souligner que notre algorithme commence par sélectionner, parmi les ancêtres inscrits dans le fichier d'une généalogie ascendante, tout individu qui y occupe plusieurs rangs, c'est-à-dire qui y apparaît plus d'une fois comme ancêtre. Car, bien entendu, si tel est le cas, il faut envisager la possibilité qu'un tel individu, appelons-le X , soit au centre d'une chaîne de consanguinité qui doit être prise en compte dans le calcul du coefficient. La vérification de cette éventualité a donc nécessité la conception de tests successifs, pierre angulaire de l'algorithme, qui d'une part certifient la possibilité que X soit bien un ancêtre commun au père et à la mère d'ego, et qui d'autre part attestent le fait que les deux ancêtres occupant les rangs immédiatement antérieurs à X ne se retrouvent pas eux-mêmes comme ancêtres communs d'une autre chaîne de consanguinité⁴.

Précisons par ailleurs que cet algorithme requiert la création de trois fichiers auxiliaires au fichier généalogique, et entraîne le traitement des variables associées au rang, au matricule et au sexe des individus.

-
3. Pour un exposé détaillé de ce coefficient, voir l'annexe.
 4. Ce faisant, nous supposons que tous les ancêtres inscrits dans une chaîne de consanguinité donnée ne sont pas déjà eux-mêmes mentionnés comme ancêtres communs dans une autre chaîne de consanguinité qui recouvrirait la première. Il peut s'ensuivre une très légère sous-estimation du coefficient, qui ne peut se rencontrer que dans des généalogies s'étendant au-delà de neuf générations.

Tableau 4

Exemple de sortie d'un calcul de coefficient de consanguinité

CALCULS DE LA CONSANGUINITÉ DE L'INDIVIDU # 2943

Nombre d'ancêtres total : 126

Nombre d'ancêtres répétés : 15

DÉTAIL DES CHAÎNES DE CONSANGUINITÉ

Matricule des ancêtres

<u>Valeur</u>	<u>Ancêtre antérieur</u>	<u>Ancêtre commun</u>	<u>Ancêtre antérieur</u>
(0,0078125)	3611- 3997- 2228-	2218-	2225- 2470- 2535
(0,0078125)	3611- 3997- 2228-	2219-	2225- 2470 2535
(0,0019531)	3611- 3997- 3987- 5411-	7092-	4282- 2484- 2532- 2535
(0,0019531)	3611- 3997- 3987- 5411-	4256-	4282- 2484- 2532- 2535

TABLEAU SYNTHÈSE DU COEFFICIENT DE CONSANGUINITÉ D'EGO

<u>Génération</u>	<u>Nombre d'ancêtres communs</u>	<u>Consanguinité partielle^a</u>	<u>Consanguinité restreinte, val. absol.^b</u>
		<u>Valeur absolue</u>	<u>% du total</u>
4	2	0,0156250	80,00
5	2	0,0039063	20,00
			0,0195313

COEFFICIENT DE CONSANGUINITÉ : 0,01953125

Sources : Registre de population de l'île de Saint-Barthélemy et logiciel OGAIP.

- Coefficient obtenu par la seule prise en compte des chaînes de consanguinité dont l'ancêtre commun est positionné en nième génération.
- Cumul des coefficients de consanguinité partielle, ou coefficient obtenu par la prise en compte des chaînes de consanguinité dont l'ancêtre commun est positionné entre la première et la nième génération.

Les résultats du calcul du coefficient de consanguinité tel que fournis par le module de génétique implanté dans OGAIP comprennent, outre la valeur même du coefficient, la liste des éléments qui ont permis de l'obtenir (voir tableau 4). Cette décomposition s'avère intéressante pour l'interprétation d'un coefficient. En effet, appliquée à l'échelle d'une communauté, par exemple, elle peut conduire à la mise en évidence d'un phénomène d'endogamie de lignée. En outre, d'un strict point de vue méthodologique, il devient possible de discerner quels sont les éléments (entre le nombre de chaînes de consanguinité et la profondeur généalogique qu'elles atteignent) qui font varier significativement la valeur du coefficient.

CONCLUSION

Le développement d'un modèle de généalogie ascendante qui puisse faire l'objet de traitements informatisés ne s'est pas limité à la seule conceptualisation de ce dernier. En effet, les exemples de fonctions d'extraction de résultats destinés à l'analyse d'ascendances illustrent la portée de la mise au point de notre modèle.

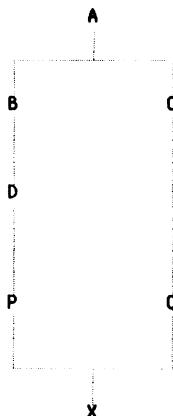
Bien que la sortie de listes et le calcul des coefficients de complétude et de consanguinité ne représentent pas des innovations de la méthodologie propre à l'approche généalogique, nous nous permettons d'insister sur les apports de leur mise en opération au sein du logiciel OGAIP. La parfaite définition (et compréhension) de ces concepts qu'a exigée leur traduction en termes d'algorithme, auront d'abord permis le calcul automatique du résultat propre à chacun. Mais c'est davantage l'importance de la prise en considération des éléments intervenant dans ces calculs qui doit être soulignée. Pensons ici aux listes qui permettent de différencier les individus inscrits dans une ascendance selon leur statut généalogique, au détail des chaînes de consanguinité comme complément du même coefficient, à la distribution des rangs d'une généalogie qui ajoute beaucoup aux coefficients de complétude. Ces données s'avéreront de précieuses informations au moment de l'analyse.

ANNEXE

Description du coefficient de consanguinité⁵

Le coefficient de consanguinité «f» de Wright exprime la probabilité qu'une paire d'allèles localisée dans les gamètes d'un individu X soit identique par descendance. La procédure consiste à retracer la généalogie de l'individu X jusqu'aux ancêtres communs de ses parents et à calculer les probabilités de transmission de l'allèle à chaque ségrégation.

La figure ci-dessous, tirée de Falconer (1960), permet d'illustrer la procédure.



Dans cet exemple, X est l'individu dont on veut calculer le coefficient de consanguinité «f». Ses parents sont P et Q. L'individu A est l'ancêtre commun à P et à Q. On veut connaître la probabilité que X reçoive deux allèles identiques transmis par P et Q en provenance de A.

Considérons B et C. La probabilité qu'ils reçoivent de A une copie d'un même gène est égale à 0,5 et la probabilité qu'ils reçoivent un gène différent est également de 0,5. S'ils reçoivent des gènes différents de A, la probabilité que ces gènes soient identiques à cause d'une consanguinité antérieure correspond au coefficient de consanguinité de A. Ainsi, la

5. Extrait de Falconer (1960 : 86-87); notre traduction.

probabilité que B et C reçoivent des gènes identiques de A est égale à $0,5(1 + f_a)$, c'est-à-dire la probabilité que deux gamètes prises au hasard chez A contiennent des allèles identiques.

Poursuivons le tracé passant par B. La probabilité que B transmette à D le gène obtenu de A égale 0,5, de D à P elle égale 0,5 et de P à X elle égale 0,5. La même chose vaut évidemment pour l'autre partie de la généalogie, celle qui passe par C et Q.

Si on met ensemble ces calculs, on trouve que la probabilité que X reçoive de A des allèles identiques est égale à

$$0,5(1 + f_a)^{3+2}$$

ou, en termes généraux, à $0,5(1 + f_a)^{n_1 + n_2}$

où n_1 est le nombre de générations à partir d'un parent de X jusqu'à l'ancêtre commun, et où n_2 est le nombre de générations à partir de l'autre parent. Si les deux parents ont plus d'un ancêtre commun, les probabilités indépendantes pour chacun des ancêtres communs doivent être additionnées pour donner le coefficient de consanguinité de la progéniture de ces parents (P et Q), d'où la formule pour calculer le coefficient de consanguinité d'un individu :

$$f_x = \sum \left[(0,5)^{n_1 + n_2 + 1} (1 + f_a) \right].$$

RÉFÉRENCES BIBLIOGRAPHIQUES

- BIDEAU, A., et M. POULAIN, 1984. «De la généalogie à la démographie historique : généalogies ascendantes et analyse démographique». Annales de démographie historique, 55-68.
- BOUCHARD, G., C. LABERGE, C.R. SCRIVER, F. GLORIEUX, M. DECLOS, L. ROBERGE, J. LAROCHELLE et S. MORTEZAI, 1984. «Étude démographique et généalogique de deux maladies héréditaires au Saguenay». Cahiers québécois de démographie, 13, 1, 118-137.
- BOUCHARD, G., R. ROY, M. DECLOS, K. KOULADJIAN et J. MATHIEU, 1988. «La diffusion du gène de la dystrophie myotonique au Saguenay (Québec)». Journal de génétique humaine, 36, 3, 221-237.
- DECLOS, M., 1985. Étude démographique d'une sous-population de la région du Saguenay-Lac-St-Jean définie par une maladie héréditaire : la tyrosinémie. Université de Montréal, Département de démographie, mémoire de maîtrise, 123 p.
- GARÇON, C. et A. JACQUARD, 1971. «Programmes de calcul des structures génétiques et des coefficients de consanguinité». In Génétique et population, Hommage à Jean Sutter, Travaux et Documents, Cahier no 60, INED, Paris, PUF, 221-238.
- FALCONER, D.S., 1960. Introduction to Quantitative Genetics. Edinburg and London, Oliver and Boyd, 365 p.
- JACQUARD, A., 1974. Génétique des populations humaines. Paris, PUF, 220 p.
- KOULADJIAN, K., 1986. Une mesure d'entropie généalogique. Chicoutimi, SOREP, Centre universitaire de recherches sur les populations, Document III-C-43, 4 p.
- KUDO, A., 1962. «A method for calculating the inbreeding coefficient». American Journal of Human Genetics, 14, 426-432.
- LÉTOURNEAU, E., 1986. Nuptialité et transmission d'une maladie héréditaire : étude de l'hypoacousie dans l'île de Saint-Barthélemy, XIXème, XXème siècles, Université de Montréal, Département de démographie, mémoire de maîtrise, 177 p.

- MAYER, F.M., C. BONAITI et J. BENOIST, 1984. «Utilisation de l'approche généalogique pour l'étude génétique de l'hypoacousie dans un isolat de la caraïbe». Anthropologie et Sociétés, 8, 2, 161-177.
- PÉLISSIER, J.P., 1983. Démographie, généalogie, micro-informatique. Société de démographie historique, Paris, 163 p.
- WRIGHT, S., 1922. «Coefficients of inbreeding and relationship». American Naturalist, 56, 330-338.

RÉSUMÉ - SUMMARY - RESUMEN

LÉTOURNEAU Esther et Francine M. MAYER - UN MODÈLE D'ANALYSE DE GÉNÉALOGIE ASCENDANTE

Cet article présente un modèle de généalogie ascendante, la définition et l'évaluation des fonctions qui permettent d'en faire l'étude, ainsi que les principaux algorithmes développés pour en faire des outils de travail sur micro-ordinateur. Les exemples de fonctions d'extraction de résultats destinés à l'analyse d'ascendances illustrent la portée de la mise au point du modèle, que ce soit pour la sortie de listes ou le calcul des coefficients de consanguinité et de complétude.

LÉTOURNEAU Esther and Francine M. MAYER - A MODEL FOR ANALYSING ASCENDING GENEALOGIES

This paper presents a model of ascending genealogies, with the definition and investigation of the functions allowing for the study of these genealogies as well as with the main algorithms used in operationalizing the model on a micro-computer. Examples of the type of results one may obtain with this model are provided; they concern file outputs as well as the computation of consanguinity and completeness coefficients.

LÉTOURNEAU Esther y Francine M. MAYER - UN MODELO DE ANÁLISIS DE GENEALOGÍA ASCENDENTE

Este artículo presenta un modelo de genealogía ascendente, la definición y la evaluación de funciones que permiten estudiar el mismo así como los principales algoritmos desarrollados para hacer de ellos herramientas de trabajo para micro-computadora. Los ejemplos de funciones de extracción de resultados destinados al análisis de ascendencias ilustran la amplitud de puesta a punto del modelo, sea por la salida de listas o por el cálculo de coeficientes de consanguinidad y de completitud.